



Amplitude onsets and spectral energy in perceptual experience

Sophie Scott* and Carolyn McGettigan

Institute of Cognitive Neuroscience, University College London, London, UK

*Correspondence: sophie.scott@ucl.ac.uk

A commentary on

A temporal sampling framework for developmental dyslexia

by Goswami, U. (2011). *Trends Cogn. Sci. (Regul. Ed.)* 15, 3–10.

In a recent review, Usha Goswami outlined a central role for disordered processing of the rise times of acoustic signals in developmental dyslexia. This approach was placed within the context of amplitude modulation spectra in sounds and in

neural oscillations, with an emphasis on problems with rise time processing at low frequency amplitude modulations (1.5–4 Hz), which are proposed to form an “edge” to syllables in speech (Goswami, 2011). In the perceptual experience of

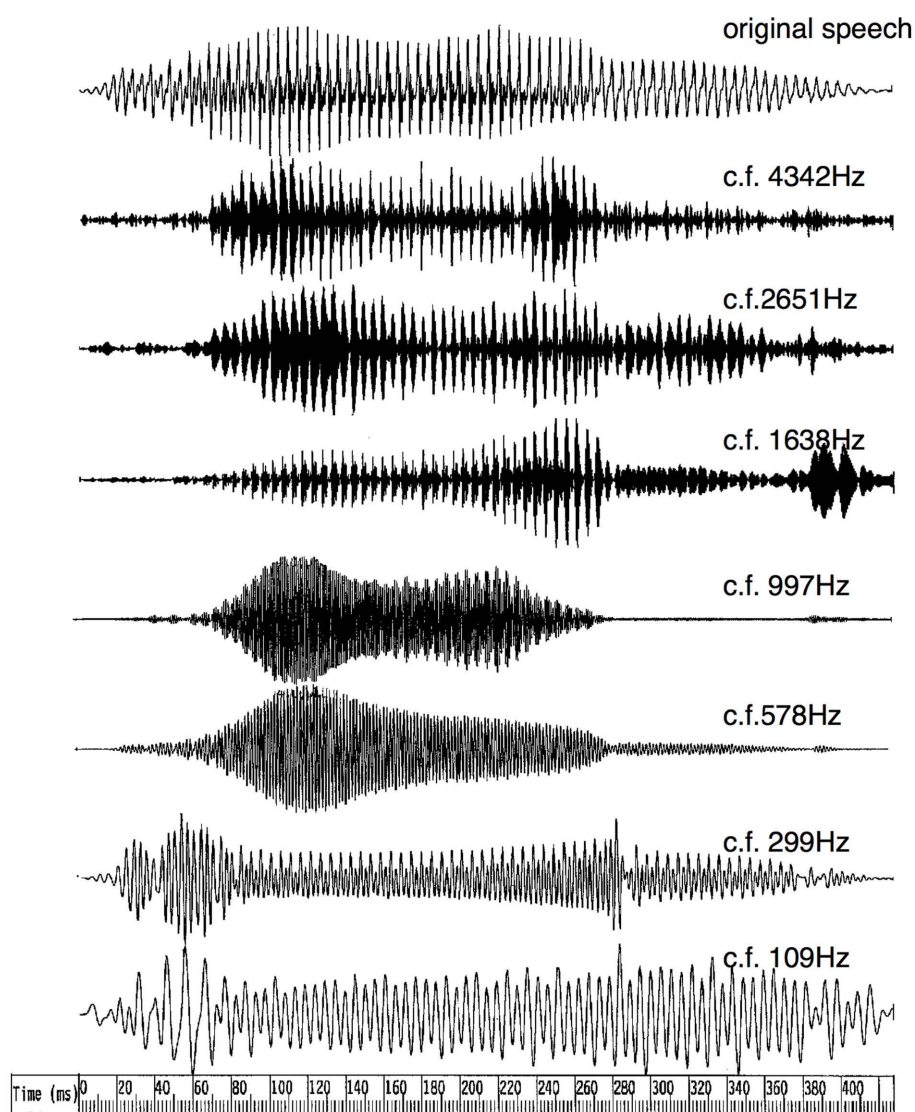


FIGURE 1 | Oscillograms of the word “one” spoken by female speaker, and the output of a gammatone filterbank when this signal is passed through (number of channels = 7, ERB = 4.0). The rise times within the different bands of spectral energy differs both in rate of change and where the big changes in rise time occur.

sound sequences, however, not all signal rise times are equal in their effects, and the onset phenomena that Goswami discusses need to be considered in the context of the spectral energy of sounds. For example, Goswami correctly stresses that people with dyslexia experience problems with onset/rhyme processing in syllables (e.g., in producing spoonerisms), but the rise time of low amplitude modulations in a syllable is not a useful marker of the onset/rhyme distinction, unless the syllable happens to start with a vowel.

In the auditory periphery, amplitude variation is represented within different channels of spectral energy (e.g., Irino and Patterson, 2001), and the rise times of the energy within these channels can vary considerably. For example, **Figure 1** shows the amplitude variation within seven spectral energy bandwidths in the spoken word “one.” Across the different spectral energy channels, rise times of the amplitude envelopes clearly vary both in the length of the rise time (how long the rise time takes), and in the point when, relative to the onset of the word, there are increases in amplitude that constitute the rise times.

The ways that different kinds of spectral energy vary in a sound have specific consequences for the rhythmic phenomena Goswami addresses. A perceptually regular sequence of speech (e.g., someone counting from 1 to 10) has no corresponding physical regularity of the onsets of the sounds:

instead, perceptual “moments of occurrence,” or perceptual centers, have been identified as the aspects of sounds – the beats – which are equally timed in both the perception and production of rhythmic speech (Morton et al., 1976). Perceptual centers are associated with increases in mid range spectral energy (around 500–1500 Hz; Marcus, 1981), i.e., with the onsets of the first formants in speech. Perceptual centers are thus linked to the onsets of vowel sounds within syllables (Cummins and Port, 1998; Scott, 1998). The perceptual center of “throw” is so much later than that of “row,” since the onsets of the more sonorous aspects of the spectral energy is much later in “throw.” These differences influence both how talkers time the utterances of “throw” and “row” when speaking rhythmically, and how listeners set those same speech items to a rhythm. Perceptual centers thus allow us to map between the ways that rhythms in sequences are both heard and produced (Morton et al., 1976).

Goswami links rise times to the “edges” of auditory objects. However, it is the perceptual center of a syllable, not its edge, which it linked with its onset and rhyme (Marcus, 1981), and it is the beat of a sound, not its edge, which drives its rhythmic properties (Terhardt and Schütte, 1976; Gordon, 1987). We agree with Goswami that it is essential to try and expand on the concept of what “phonological” problems seen in dyslexia truly entail in acoustic

terms, and we suggest that the perceptual centers of sounds can capture the properties of auditory objects that rise times alone cannot.

REFERENCES

- Cummins, F., and Port, R. (1998). Rhythmic constraints on stress timing in English. *J. Phon.* 26, 145–171.
- Gordon, J. W. (1987). The perceptual attack time of musical tones. *J. Acoust. Soc. Am.* 82, 88–105.
- Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci. (Regul. Ed.)* 15, 3–10.
- Irino, T., and Patterson, R. D. (2001). A compressive gam-machirp auditory filter for both physiological and psychophysical data. *J. Acoust. Soc. Am.* 109, 2008–2022.
- Marcus, S. M. (1981). Acoustic determinants of perceptual centre (P center) location. *Percept. Psychophys.* 30, 247–256.
- Morton, J., Marcus, S. M., and Frankish, C. (1976). Perceptual centres (P centers). *Psychol. Rev.* 83, 405–408.
- Scott, S. K. (1998). The point of P-centres. *Psychol. Res.* 61, 4–11.
- Terhardt, E., and Schütte, H. (1976). Akustische Rhythmus-Wahrnehmung: Subjektive Gleichmässigkeit. *Acustica* 35, 122–126.

Received: 19 February 2012; accepted: 01 March 2012; published online: 27 March 2012.

Citation: Scott S and McGettigan C (2012) Amplitude onsets and spectral energy in perceptual experience. *Front. Psychology* 3:80. doi: 10.3389/fpsyg.2012.00080
This article was submitted to *Frontiers in Auditory Cognitive Neuroscience*, a specialty of *Frontiers in Psychology*.
Copyright © 2012 Scott and McGettigan. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.